GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM

CROSS REFERENCE TO RELATED CASES

This is a division of co-pending Application No. 09/746,801, filed December 20, 2000, which is a continuation-in-part of Application No. 09/513,057, filed February 24, 2000, now Patent No. 6,433,251, which is a continuation-in-part of International Application No. PCT/US99/18747, filed August 17, 1999, which claims the benefit of Provisional Application No. 60/096,802, filed August 17, 1998. All of these applications are incorporated herein in their entirety.

FIELD OF THE INVENTION

This invention relates to genes that regulate circadian clock functions and photoperiodism in plants, and relates in particular to the *ELF3* gene. Aspects of the invention include the purified *ELF3* gene product (ELF3 protein), as well as nucleic acid molecules encoding this gene product. Nucleic acid vectors, transgenic cells, and transgenic plants having modified ELF3 activity are also provided.

BACKGROUND OF THE INVENTION

Shoot development in flowering plants is a continuous process ultimately controlled by the activity of the shoot apical meristem. Apical meristem activity during normal plant development is sequential and progressive, and can be summarized as a series of overlapping phases: vegetative→inflorescence→floral (V→I→F). Over the past 50 years many models have been proposed for the control of the vegetative-to-floral transition. These models range from simple single pathway models to complex multiple pathway models, and are largely based on physiological studies (for review, see Bernier, 1988). Modern techniques provide researchers with genetic and molecular methods that can be used to further investigate the control of V→I→F transitions.

One such modern technique now routinely practiced by plant molecular biologists is the production of transgenic plants carrying a heterologous gene sequence. Methods for incorporating an isolated gene sequence into an expression cassette, producing plant transformation vectors, and transforming many types of plants are well known. Examples of the production of transgenic plants having modified characteristics as a result of the introduction of a heterologous transgene include: U.S. Patent Nos. 5,268,526 (modification of phytochrome expression in transgenic plants); 5,719,046 (production of herbicide resistant plants by introduction of bacterial dihydropteroate synthase gene); 5,231,020 (modification of flavenoids in plants); 5,583,021 (production of virus resistant plants); and 5,767,372 and 5,500,365 (production of insect resistant plants by introducing *Bacillus thuringiensis* genes).

Light quality, photoperiod, and temperature often act as important, and for some species essential, environmental cues for the initiation of flowering. However, there is very little information on the molecular mechanisms that directly regulate the developmental pathway from reception of the inductive light signal(s) to the onset of flowering and the initiation of floral meristems. The analysis of floral transition mutants in pea *(Pisum sativum)* (see Murfet, 1985) and *Arabidopsis* (see Koornneef *et al.*, 1991) has demonstrated that at least part of the genetic hierarchy controlling flowering onset is responsive to the number of hours of light perceived by a plant within a 24 hour light/dark cycle. The monitoring of the length of the light period is referred to as the photoperiodic response. Photoperiodic responses have long been thought to be tied to one or more biological clocks that regulate many physiological and developmental processes on the basis of an endogenous circadian rhythm.

Many important physiological and developmental plant processes are influenced by circadian rhythms. These include the induction of gene transcription, leaf movement, stomatal opening, and the photoperiodic control of flowering. While the relationship of these plant processes to the circadian rhythm has long been recognized, the genetic analysis of circadian rhythms in plants has only recently begun. Most of the genetic analysis of circadian regulation has been performed with *Drosophila* and *Neurospora crassa*, where mutational studies have led to the isolation of the *per* and *frq* genes, respectively (Hall, 1990; Dunlap, 1993). These genes are thought to encode components of the circadian oscillator, in part because, while null alleles cause arrhythmic responses, alleles of these genes exist that produce either long or short period responses. Transcriptional production of *per* and *frq* mRNA cycles on a twenty-four hour period, and both genes regulate their own expression (Edery *et al.*, 1994; Aronson *et al.*, 1994).

Arabidopsis is a quantitative long-day (LD) plant — wild-type plants will initiate flowering more quickly when grown under LD light conditions than when grown under short-day (SD) light conditions. In order to identify genes required for floral initiation and development, populations of *Arabidopsis thaliana* ecotype Columbia grown in SD conditions have been screened for early-flowering mutants. Isolated mutants were then examined for additional shoot development anomalies, and those with discreet shoot phenotypes related to meristem function or light perception were considered for further analysis. Such mutants may identify genes that are part of functionally redundant pathways that operate, to varying degrees, as "fail-safe" mechanisms for ensuring shoot growth and reproductive development. Examples of such functionally redundant pathways have been described in studies of *Drosophila* (*e.g.*, Hülskamp *et al.*, 1990) and *C. elegans* (*e.g.*, Lambie and Kimble, 1991). The key genes identified by these *Arabidopsis* screens were the *TERMINAL FLOWER* 1 (*TFL1*) gene and the *EARLY-FLOWERING* 3 (*ELF3*) gene (Shannon and Meeks-Wagner, 1991; Zagotta *et al.*, 1992).

The *early-flowering* (*elf3*) mutant of *Arabidopsis* is insensitive to photoperiod with regard to floral initiation. Plants homozygous for a mutation in the *ELF3* locus flower at the same time in LD

and SD growth conditions, whereas floral initiation of wild-type plants is promoted by LD growth conditions (Zagotta *et al.*, 1992; Zagotta *et al.*, 1996). In LD conditions, the flowering time of the *elf3-1* heterozygote is intermediate between wild-type and the homozygous mutant. In addition to being photoperiod-insensitive, all *elf3* mutants display the long hypocotyl phenotype characteristic of plants defective in light reception or the transduction of light signals (Zagotta *et al.*, 1992; Zagotta *et al.*, 1996). The majority of long hypocotyl mutants that have been identified are defective in red light-mediated inhibition of hypocotyl elongation. In contrast, *elf3* mutants are primarily defective in blue light-dependent inhibition of hypocotyl elongation, although they are also partially deficient in red light-dependent inhibition of hypocotyl elongation (Zagotta *et al.*, 1996).

The availability of the *ELF3* gene would facilitate the production of transgenic plants having altered circadian clock function and programmed photoperiodic responses. It is to such a gene that the present invention is directed.

SUMMARY OF THE INVENTION

The invention provides an isolated *ELF3* gene from *Arabidopsis* that is shown to complement the *elf3* photoperiod-insensitive flowering and elongated hypocotyl defects when introduced into *elf3* mutant plants.

One aspect of this invention is a purified protein having ELF3 protein biological activity. The prototypical *Arabidopsis* ELF3 protein has the amino acid sequence shown in SEQ ID NO: 2. Variants of this protein that differ from SEQ ID NO: 2 by one or more conservative amino acid substitutions are also provided, as are homologs of the ELF3 protein. Such homologs typically share at least 60% sequence identity with the sequence shown in SEQ ID NO: 2. Nucleic acid molecules encoding these proteins are also part of this invention. Such nucleic acid molecules include those having the nucleotide sequences set forth in SEQ ID NO: 1, SEQ ID NO: 3, and SEQ ID NO:4.

Recombinant nucleic acid molecules in which a promoter sequence is operably linked to any of these ELF3 protein-encoding nucleic acid sequences are further aspects of this invention. The invention also provides cells transformed with such a recombinant nucleic acid molecule and transgenic plants comprising the recombinant nucleic acid molecule. Such transgenic plants may be, for instance, *Arabidopsis*, pepper, tomato, tobacco, broccoli, cauliflower, cabbage, canola, bean, soybean, rice, corn, wheat, barley, citrus, cotton, cassava and walnut, trees such as poplar, oak, maple, pine, spruce, and other conifers, and ornamental plants (*e.g.*, petunias, orchids, carnations, roses, impatiens, pansies, lilies, snapdragons, geraniums, and so forth).

A further aspect of this invention is an isolated nucleic acid molecule or oligonucleotide comprising 15, 20, 30, 50, or 100 contiguous nucleotides of the sequence shown in SEQ ID NOs: 1, 3, or 4. Such nucleic acid molecules or oligonucleotides may be operably linked to a promoter sequence, and

may be in the sense or antisense orientation in relation to such a promoter. The invention also includes cells and plants transformed with such recombinant nucleic acid molecules, with or without an attached promoter.

Further embodiments of this invention include isolated nucleic acid molecules that hybridize under specified hybridization conditions to the nucleic acid sequence set forth in SEQ ID NO: 1, and that encode a protein having ELF3 protein biological activity. Closely related *ELF3* gene homologs may be detected by hybridization under stringent conditions, whereas less closely related homologs may be detected by hybridization at low stringency. Appropriate wash conditions for stringent hybridization may be 55° C, 0.2 x SSC and 0.1% SDS for 1 hour. Appropriate wash conditions for low stringency hybridization may be 50° C, 2 x SSC, 0.1 % for 3 hours. Such a hybridizing isolated nucleic acid molecule may be operably linked to a promoter for expression in plants. Cells transformed with such a recombinant nucleic acid molecule, and transgenic plants that comprise such a molecule, are also provided.

The invention also provides the 5' regulatory region of the *ELF3* gene. This regulatory region, or parts thereof, may be used to obtain ELF3-like circadian-rhythm expression of particular genes. For example, the *ELF3* 5' regulatory region may be operably linked to an open reading frame of a gene of interest, and the resulting recombinant construct may be introduced into a plant by transformation. One embodiment of an ELF3 regulatory region is about nucleotides 1 through about 1900 of the 5' upstream region shown in SEQ ID NO: 5.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. Sequence comparison of *ELF3* homologs.

Multiple-sequence alignment of ELF3 (residues 1-695 of SEQ ID NO: 2) and several putative ELF3 homologs from *Arabidopsis thaliana* (*Essence of ELF3 Consensus, EEC*) (residues 1-540 of SEQ ID NO: 33) and other plant species (*Cardamine oligosperma* (residues 1-577 of SEQ ID NO: 13), tomato (residues 1-179 of SEQ ID NO. 24 and residues 1-389 of SEQ ID NO: 23), rice (residues 1-760 of SEQ ID NO: 27), and maize (residues 117-247 of SEQ ID NO: 29)). Protein designations are given on the left in the same order. Amino acid residues are numbered on the right. Residues shaded in black indicate identity of at least three ELF3/ELF3-related sequences in the alignment; light-shaded residues indicate similarity to consensus. Nucleotide sequences from *C. oligosperma* (a member of the family Brassicaceae) were obtained by sequencing polymerase chain reaction products using degenerate oligos to the *Arabidopsis ELF3* gene and genomic DNA or cDNA prepared from *C. oligosperma* seedlings. Sequences were aligned and analyzed using CLUSTAL W (J. D. Thompson, D. G. Higgins, T. Gibson, *Nucleic Acids Res.* **22**, 4673-80, 1994) and PrettyBox (Genetics Computer Group, Inc., Madison, WI).

Figure 2. Sequence comparison of *ELF3* homologs showing consensus boxes.

Multiple-sequence alignment shows four highly conserved regions within ELF3 and putative ELF3 homologs from *Arabidopsis thaliana* (*Essence of ELF3 Consensus, EEC*) and other plant species (*Cardamine oligosperma*, tomato, rice, and maize). Protein designations are given on the left in the same order. Amino acid residues are numbered on both the right and left. Residues shaded in black indicate identity of at least three ELF3/ELF3-related sequences in the alignment; light-shaded residues indicate similarity to consensus. Sequences were aligned and analyzed using CLUSTAL W (J. D. Thompson, D. G. Higgins, T. Gibson, *Nucleic Acids Res.* 22, 4673-80, 1994) and PrettyBox (Genetics Computer Group, Inc., Madison, WI).

GenBank accession numbers for *ELF3* and putative *ELF3* homologs are as follows: *AtELF3* (*A. thaliana* genomic DNA: AC004747, published December 17, 1999), *AtEEC* (*A. thaliana* genomic DNA: AB023045, published November 20, 1999), *cELF3* (yet to be submitted), *tELF3* [*Lycopersicon esculentum* Expressed Sequence Tags (ESTs) from Clemson University Genomics Institute: AW093790 (October 18, 1999), AI894513 (July 27, 1999), AI488927 (June 29, 1999), AI486934 (June 29, 1999), AI894398 (July 27, 1999)], *rELF3* (*Oryza sativa* genomic DNA: AP000399, published December 3, 1999), *mELF3* (*Zea mays* EST from Stanford University Genome Center: AI637184, published April 26, 1999).

In Block I, the "AtELF3" amino acid sequence corresponds to residues 13-49 of SEQ ID NO: 2; the "AtEEC" amino acid sequence corresponds to residues 15-51 of SEQ ID NO: 33; the "cardamineELF3" amino acid sequence corresponds to residues 13-49 of SEQ ID NO: 13; the "tomatoELF3" amino acid sequence corresponds to residues 13-49 of SEQ ID NO: 24; and the "riceELF3" amino acid sequence corresponds to residues 22-59 of SEQ ID NO: 27.

In Block II, the "AtELF3" amino acid sequence corresponds to residues 317-365 of SEQ ID NO: 2; the "AtEEC" amino acid sequence corresponds to residues 238-286 of SEQ ID NO: 33; the "cELF3" amino acid sequence corresponds to residues 291-339 of SEQ ID NO: 13; the "tELF3" amino acid sequence corresponds to residues 22-70 of SEQ ID NO: 23; the "rELF3" amino acid sequence corresponds to residues 394-442 of SEQ ID NO: 27; and the "maizeELF3" amino acid sequence corresponds to residues 22-70 of SEQ ID NO: 57.

In Block III, the "AtELF3" amino acid sequence corresponds to residues 462-486 of SEQ ID NO: 2; the "AtEEC" amino acid sequence corresponds to residues 358-379 of SEQ ID NO: 33; the "cELF3" amino acid sequence corresponds to residues 441-464 of SEQ ID NO: 13; the "tELF3" amino acid sequence corresponds to residues 167-189 of SEQ ID NO: 23; the "rELF3" amino acid sequence corresponds to residues 544-565 of SEQ ID NO: 27; and the "mELF3" amino acid sequence corresponds to residues 162-178 of SEQ ID NO: 57.

In Block IV, the "AtELF3" amino acid sequence corresponds to residues 660-687 of SEQ ID NO: 2; the "AtEEC" amino acid sequence corresponds to residues 505-532 of SEQ ID NO: 33; the "cELF3" amino acid sequence corresponds to residues 639-653 of SEQ ID NO: 14; the "tELF3" amino acid sequence corresponds to residues 358-385 of SEQ ID NO: 23; the "rELF3" amino acid sequence corresponds to residues 729-756 of SEQ ID NO: 27; and the "mELF3" amino acid sequence corresponds to residues 285-312 of SEQ ID NO: 57.

Figure 3 is a Table showing growth and flowering characteristics of *Arabidopsis* seedlings over-expressing ELF3 (ELF3-OX), seedlings that are mutant in ELF3 (*elf3-1*).

Figure 4 shows the features of the predicted 695 amino acid ELF3 protein, and the molecular basis of the several *elf3* mutations.

SEQUENCE LISTING

The nucleic and amino acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases, and three letter code for amino acids. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand.

SEQ ID NO: 1 shows the cDNA and amino acid sequence of *Arabidopsis ELF3*.

SEQ ID NO: 2 shows the amino acid sequence of *Arabidopsis* ELF3 protein.

SEQ ID NO: 3 shows the genomic sequence of *Arabidopsis ELF3*. The sequence comprises the following regions:

Nucleotides	Feature
1-142	promoter region
143-424	exon 1 (5' UTR)
425-644	exon 1 continued (initiating ATG at 425)
645-1006	intron 1
1007-1803	exon 2
1804-2983	intron 2
2984-3037	exon 3
3038-3127	intron 3
3128-4142	exon 4
4143-4145	stop codon

4146-4221 3' UTR and 3' regulatory region.

SEQ ID NO: 4 shows the DNA and corresponding amino acid sequence of the *Arabidopsis ELF3* ORF.

SEQ ID NO: 5 shows the 4071 base pair *Arabidopsis ELF3* 5' regulatory region.

SEQ ID NO: 6-11 show primers that can be used to amplify certain portions of the *Arabidopsis ELF3* sequence.

SEQ ID NO: 12 shows the cDNA and corresponding amino acid sequence of the *Cardamine oligosperma* ELF3 ortholog, cELF3. This sequence can also be determined by applying well known computer analyses to the genomic sequence shown in SEQ ID NO: 14 (also referred to as COELF3 ~ 1) to determine where the introns and exons are.

SEQ ID NO: 13 (also referred to as COELF3 ~ 2) shows the amino acid sequence of the *Cardamine oligosperma* ELF3 ortholog, cELF3.

SEQ ID NO: 14 (also referred to as COELF3 ~ 1) shows the genomic sequence of the *Cardamine oligosperma* ELF3 ortholog, cELF3.

SEQ ID NO: 15 shows a partial DNA sequence (also referred to as PEAELF~2) of the pea ELF 3 ortholog.

SEQ ID NO: 16 (also referred to as PEAELF~1) shows the amino acid sequence of the partial pea ELF 3 ortholog.

SEQ ID NO: 17 (also referred to as BROCCA~2) shows the amino acid sequence of the broccoli/cauliflower EEC protein.

SEQ ID NO: 18 shows a partial DNA (also referred to as GMELF3~2) sequence of the *Glycine max* (soybean) ELF3 coding region.

SEQ ID NO: 19 (also referred to as GMELF3~1) shows the amino acid sequence of the partial *Glycine max* (soybean) ELF3 protein.

SEQ ID NO: 20 shows the DNA (also referred to as BROCCA~1) a sequence of the *Lycopersicon esculentum* (tomato) ELF3 (N-terminus #2) coding region.

SEQ ID NO: 21 shows the DNA (also referred to as LEAFFO~1) sequence of the *Lycopersicon esculentum* (tomato) ELF3 (N-terminus #1) coding region.

SEQ ID NO: 22 shows the DNA (also referred to as LE5B39~1) sequence of the *Lycopersicon esculentum* (tomato) coding region.

SEQ ID NO: 23 (also referred to as LEELF3~3) shows the amino acid sequence of the *Lycopersicon esculentum* (tomato) ELF3 (C-terminus) coding region.

SEQ ID NO: 24 (also referred to as LEELF~2) shows a partial amino acid sequence of the *Lycopersicon esculentum* (tomato) protein.

SEQ ID NO: 25 (also referred to as LEELF3~1) shows the amino acid sequence of the *Lycopersicon esculentum* (tomato) ELF3 (N-terminus #2) protein.

SEQ ID NO: 26 shows the DNA (also referred to as OSELF3~2) sequence of the *Oryza sativa* (rice) ELF3 genomic region.

SEQ ID NO: 27 (also referred to as OSELF3~1) shows the amino acid sequence of the *Oryza sativa* (rice) ELF3 protein.

SEQ ID NO: 28 shows a partial DNA (also referred to as ZM8CC4~1) sequence of the *Zea mays* (maize) ELF3 coding region.

SEQ ID NO: 29 (also referred to as ZMELF3~2) shows the amino acid sequence of the partial *Zea mays* (maize) ELF3 protein.

SEQ ID NO: 30 shows a partial DNA (also referred to as ZMELF3~4) sequence of the *Zea mays* (maize) ELF3 #2 coding region.

SEQ ID NO: 31 (also referred to as ZMELF3~3) shows the amino acid sequence of the partial *Zea mays* (maize) ELF3 #2 coding region.

SEQ ID NO: 32 shows the DNA (also known as ATEECG~1) of the *Arabidopsis thaliana* EEC genomic region.

SEQ ID NO: 33 (also known as ATEECP~1) shows the amino acid sequence of the *Arabidopsis thaliana* EEC protein.

SEQ ID NO: 34 shows the DNA (also known as ATELF3~1) sequence of the *Arabidopsis thaliana* ELF3 genomic region.

SEQ ID NO: 35 (also known as ATELF3~2) shows the amino acid sequence of the *Arabidopsis thaliana* ELF3 protein.

SEQ ID NO: 36 (also known as MTELF3N1) shows a portion of exon 1, including 5'UTR and start codon, of the *Medicago trunculata* ELF3 cDNA nucleotide sequence. This partial sequence was originally reported in Genbank Accession No. AW690413.

SEQ ID NO: 37 (also known as MTELF3P1) shows the peptide portion of the *Medicago trunculata* ELF3 protein encoded for by SEQ ID NO: 36.

SEQ ID NO: 38 (also known as MTELF3N4) shows a portion of exon 4, including stop codon and 3'UTR, of the *Medicago trunculata* ELF3 nucleotide sequence. This partial sequence was originally reported as Genbank Accession No. AW693560.

SEQ ID NO: 39 (also known as MTELF3P4) shows the peptide portion of the *Medicago trunculata* ELF3 protein encoded for by SEQ ID NO: 38.

SEQ ID NO: 40 (also known as PSELF3N3) shows a portion of exon 3 to exon 4 of the *Pisum sativa* genomic DNA encoding ELF3.

SEQ ID NO: 41 (also known as PSELF3P3) shows the peptide portion of the *Pisum sativa* ELF3 protein encoded for by SEQ ID NO:40.

SEQ ID NO: 42 (also known as PSELF3N4) shows a portion of exon 4 of the *Pisum sativa* genomic DNA encoding ELF3.

SEQ ID NO: 43 (also known as PSELF3P4) shows the peptide portion of the *Pisum sativa* ELF3 protein encoded for by SEQ ID NO: 42.

SEQ ID NO: 44 (also known as GMELF3N) shows a portion of the *Glycine max* cDNA encoding ELF3. This partial sequence was originally reported in Genbank Accession No. AW757137.

SEQ ID NO: 45 (also known as GMELF3P) shows the peptide portion of the *Glycine max* ELF3 protein encoded for by SEQ ID NO: 44.

SEQ ID NO: 46 (also known as XELF3N1) shows a portion of the *Xanthium* genomic DNA (from exon 3 to exon 4) encoding ELF3.

SEQ ID NO: 47 (also known as XELF3P1) shows the peptide portion of the *Xanthium* ELF3 protein encoded for by SEQ ID NO: 46.

SEQ ID NO: 48 (also known as XELF3N2) shows a portion of the *Xanthium* genomic DNA (from exon 3 to exon 4) encoding ELF3.

SEQ ID NO: 49 (also known as XELF3P2) shows the peptide portion of the *Xanthium* ELF3 protein encoded for by SEQ ID NO: 48.

SEQ ID NO: 50 (also known as XELF3N4) shows a portion of the *Xanthium* genomic DNA (a portion of exon 4) encoding ELF3.

SEQ ID NO: 51 (also known as XELF3P4) shows the peptide portion of the *Xanthium* ELF3 protein encoded for by SEQ ID NO: 50.

SEQ ID NO: 52 (also known as POPELF3N) shows a portion of the Poplar genomic DNA (a portion of exon 4) encoding ELF3.

SEQ ID NO: 53 (also known as POPELF3P) shows the peptide portion of the Poplar ELF3 protein encoded for by SEQ ID NO: 52.

SEQ ID NO: 54 (also known as MIMELF3N) shows a portion of the *Mimulus* genomic DNA (a portion of exon 4) encoding ELF3.

SEQ ID NO: 55 (also known as MIMELF3P) shows the peptide portion of the *Mimulus* ELF3 protein encoded for by SEQ ID NO: 54.

SEQ ID NO: 56 (also known as ZMELF3N) shows a portion of the *Zea mays* contig of cDNA/genomic DNA (exon 2, exon 3, intronic sequence, and exon 4, including stop codon and 3'UTR) encoding ELF3. This partial sequence was originally reported in Genbank Accession No. AI637184.

SEQ ID NO: 57 (also known as ZMELF3P) shows the peptide portion of the *Zea mays* ELF3 protein encoded for by SEQ ID NO: 56.

SEQ ID NO: 58 (also known as LEELF3-AN) shows a portion of the *Lycopersicon esculentum* cDNA (exon 1, exon 2, exon 3, and exon 4, including stop codon and 3'UTR) encoding ELF3.

SEQ ID NO: 59 (also known as LEELF3-AP) shows the peptide portion of the *Lycopersicon esculentum* ELF3 protein encoded for by SEQ ID NO: 58.

SEQ ID NO: 60 (also known as BRELF3AN) shows a portion of the Broccoli genomic DNA (portion of exon 1, exon 2, exon 3, and portion of exon 4) encoding ELF3.

SEQ ID NO: 61 (also known as BRELF3AP) shows the peptide portion of the Broccoli ELF3 protein encoded for by SEQ ID NO: 60.

SEQ ID NO: 62 (also known as BRELF3BN) shows a portion of the Broccoli genomic DNA (a portion of exon 4) encoding ELF3.

SEQ ID NO: 63 (also known as BRELF3BP) shows the peptide portion of the Broccoli ELF3 protein encoded for by SEQ ID NO: 62.

SEQ ID NOs: 64-68 (also known as C-FWD, D-REV, B-FWD,Pea1b-C-FWD, and C-REV, respectively) show primers used to amplify *ELF3* homologous sequences.

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

Unless otherwise noted, technical terms are used according to conventional usage. Definitions of common terms in molecular biology may be found in Benjamin Lewin, *Genes V*, published by Oxford University Press, 1994 (ISBN 0-19-854287-9); Kendrew *et al.* (eds.), *The Encyclopedia of Molecular Biology*, published by Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); and Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

In order to facilitate review of the various embodiments of the invention, the following definitions of terms are provided:

***ELF3* gene/*ELF3* cDNA:** Nucleic acid molecules that encode an ELF3 protein. Nucleic acid molecules that encode the *Arabidopsis* ELF3 protein are provided in SEQ ID NO: 3 (*Arabidopsis ELF3* gene), SEQ ID NO: 1 (*Arabidopsis ELF3* cDNA) and SEQ ID NO:4 (*Arabidopsis ELF3* open reading frame). The invention includes not only the nucleic acid molecules provided in SEQ ID NOS: 1, 3 and 4, but also homologs and orthologs of these sequences, other nucleic acid molecules that encode ELF3 proteins, and probes and primers that are derived from these sequences.

***elf3* mutant**: The *early-flowering* (*elf3*) mutant of *Arabidopsis* is insensitive to photoperiod with regard to floral initiation (Zagotta *et al.,* 1992; Zagotta *et al.*, 1996). In addition to being photoperiod-insensitive, all *Arabidopsis efl3* mutants display the long-hypocotyl phenotype characteristic of plants defective in light reception or the transduction of light signals (Zagotta *et al.,* 1992; Zagotta *et al.*, 1996). *Elf3* mutants are primarily defective in blue light-dependent inhibition of hypocotyl elongation, although *elf3* mutants are also partially deficient in red light-dependent inhibition of hypocotyl elongation (Zagotta *et al.*, 1996).

ELF3 protein: A protein having ELF3 protein biological activity and sharing amino acid sequence identity with the amino acid sequence of the prototypical ELF3 protein shown in SEQ ID NO: 2 (the *Arabidopsis* ELF3 protein). ELF3 proteins that are more distantly related to the prototypical ELF3 protein will share at least 60% amino acid sequence identity with the sequence shown in SEQ ID NO: 2, as determined by the methods described below. More closely related ELF3 proteins may share at least 70%, 75% or 80% sequence identity with the *Arabidopsis* ELF3 protein. ELF3 proteins that are most closely related to the *Arabidopsis* protein will have ELF3 protein biological activity and share at least 85%, 90% or 95% sequence identity with the *Arabidopsis* protein.

ELF3 protein biological activity: The ability of a protein to complement an *elf3* mutant. The ability of a protein to complement an *elf3* mutant may be readily determined by introducing the gene encoding the protein into an *elf3* mutant plant using standard methods. If the encoded protein has *ELF3* protein biological activity, this will be manifested as a proportion of the transgenic progeny plants having a wild-type phenotype for those characteristics linked to the *elf3* mutant (*e.g.*, photoperiod-insensitive flowering and elongated hypocotyl).

***ELF3* promoter:** The region of nucleic acid sequence upstream (5') of the *ELF3* coding sequence that is responsible for spatial and temporal regulation of *ELF3* transcription. *ELF3* transcription is circadian regulated, but with an RNA maximum that is "later" in the 24-hour period than that of other known circadian genes, *e.g.*, CAB, CCR2, CCA1 and LHY (Wang and Tobin, 1998; Schaffer *et al.*, 1998). *ELF3*-like circadian rhythm or cyclic transcriptional regulation refers to this type of a relatively delayed transcription maximum. Because *ELF3* transcription reaches a maximal level relatively late in the 24-hour period, the *ELF3* promoter will allow for altering the setting of the circadian clock. For instance, if another circadian-regulated gene (*e.g.*, *chlorophyll a/b* binding protein)

is expressed from the *ELF3* promoter, the circadian set on this protein will be delayed to match that of *ELF3*. In addition, the *ELF3* promoter may be used to provide altered expression of other genes that are under control of the circadian clock, if clock components and/or regulators such as CCA1 and LHY are driven by the *ELF3* promoter instead of their own promoters or a constitutive promoter, for instance the 35S promoter.

The *ELF3* promoter region is contained within the 4071 kb 5' regulatory region sequence shown in SEQ ID NO: 5, but one of ordinary skill in the art will appreciate that expression may be controlled by using less than this entire 5' upstream region, *e.g.*, nucleotides 500-4071, 1000-4071, 1500-4071, 2000-4071, 2500-4071, 3000-4071, 3500-4071 or 4000-4071. One embodiment of an ELF3 promoter is about nucleotides 1 through about 1900 of the 5' upstream region shown in SEQ ID NO: 5.

Sequences as short as 50 or 100 nucleotides from within the 5' regulatory region may also be employed. The degree to which such a sequence provides for *ELF3*-like circadian cyclic transcriptional regulation, when included in an expression vector, can be ascertained by the methods described herein. Thus, the term "biologically active *ELF3* promoter" refers to a 5' regulatory region of an *ELF3* gene, or a part or a variant of such a region, that, when operably linked to the 5' end of an ORF and introduced into a plant, results in ELF3-like (*i.e.*, relatively late) circadian cyclic transcript expression of the protein encoded by the ORF.

Essence of ELF3 Consensus (EEC): One or more highly conserved regions of amino acid sequence within an ELF3 protein or ELF3 protein homolog. EECs are depicted in Figures 1 and 2.

Oligonucleotide: A linear polynucleotide sequence of up to about 100 nucleotide bases in length.

Probes and primers: Nucleic acid probes and primers can be readily prepared based on the nucleic acid molecules provided in this invention. A probe comprises an isolated nucleic acid attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. Methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed, *e.g.*, in Sambrook *et al.* (1989) and Ausubel *et al.* (1987).

Primers are short nucleic acid molecules, typically DNA oligonucleotides 15 nucleotides or more in length. Primers can be annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, and then extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification of a nucleic acid sequence, *e.g.*, by the polymerase chain reaction (PCR) or other nucleic-acid amplification methods known in the art.

Methods for preparing and using probes and primers are described, for example, in Sambrook *et al.* (1989), Ausubel *et al.* (1987), and Innis *et al.* (1990). PCR primer pairs can be derived from a known

sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, © 1991, Whitehead Institute for Biomedical Research, Cambridge, MA). One of ordinary skill in the art will appreciate that the specificity of a particular probe or primer increases with its length. Thus, for example, a primer comprising 20 consecutive nucleotides of the *Arabidopsis ELF3* cDNA or gene will anneal to a target sequence such as an *ELF3* gene homolog from tomato contained within a tomato genomic DNA library with a higher specificity than a corresponding primer of only 15 nucleotides. Thus, in order to obtain greater specificity, probes and primers can be selected that comprise 20, 25, 30, 35, 40, 50 or more consecutive nucleotides of the *Arabidopsis ELF3* cDNA or gene sequences.

The invention thus includes isolated nucleic acid molecules that comprise specified lengths of the disclosed *ELF3* cDNA or gene sequences. Such molecules may comprise at least 20, 25, 30, 35, 40, 50 or 100 consecutive nucleotides of these sequences and may be obtained from any region of the disclosed sequences. By way of example, the *Arabidopsis ELF3* cDNA, ORF and gene sequences may be apportioned into halves or quarters based on sequence length, and the isolated nucleic acid molecules may be derived from the first or second halves of the molecules, or any of the four quarters. The *Arabidopsis ELF3* cDNA, shown in SEQ ID NO: 1, can be used to illustrate this. The *Arabidopsis ELF3* cDNA is 2518 nucleotides in length and so may be hypothetically divided into about halves (nucleotides 1-1259 and 1260-2518) or about quarters (nucleotides 1-629, 630-1259, 1260-1889 and 1890-2518). Nucleic acid molecules may be selected that comprise at least 20, 25, 30, 35, 40, 50 or 100 consecutive nucleotides of any of these or other portions of the *Arabidopsis ELF3* cDNA. Thus, representative nucleic acid molecules might comprise at least 25 consecutive nucleotides of the region comprising nucleotides 1-1259 of the disclosed *Arabidopsis* cDNA, or of the regions comprising nucleotides 1-1135 or 2502-2518 of the cDNA.

Sequence identity: The similarity between two nucleic acid sequences, or two amino acid sequences is expressed in terms of the similarity between the sequences, otherwise referred to as sequence identity. Sequence identity is frequently measured in terms of percentage identity (or similarity or homology); the higher the percentage, the more similar the two sequences are. Homologs of the *Arabidopsis* ELF3 protein will possess a relatively high degree of sequence identity when aligned using standard methods. Methods of alignment of sequences for comparison are well known in the art. Various programs and alignment algorithms are described in: Smith & Waterman *Adv. Appl. Math.* 2: 482, 1981; Needleman & Wunsch *J. Mol. Biol.* 48: 443, 1970; Pearson & Lipman *Proc. Natl. Acad. Sci. USA* 85: 2444, 1988; Higgins & Sharp *Gene,* 73: 237-244, 1988; Higgins & Sharp *CABIOS* 5: 151-153, 1989; Corpet *et al. Nuc. Acids Res.* 16, 10881-90, 1988; Huang *et al. Computer Appls. in the Biosciences* 8, 155-65, 1992; and Pearson *et al. Meth. Mol. Bio.* 24, 307-31, 1994. Altschul *et al.* (*J. Mol. Biol.* 215:403-410, 1990), presents a detailed consideration of sequence alignment methods and homology calculations.

The NCBI Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1990) is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, MD) and on the Internet, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastn and tblastx. It can be accessed at the NCBI BLAST web-site. A description of how to determine sequence identity using this program is available at the help page of the NCBI web-site.

Homologs of the disclosed *Arabidopsis* ELF3 protein typically possess at least 60% sequence identity counted over full length alignment with the amino acid sequence of *Arabidopsis* ELF3 using the NCBI Blast 2.0, gapped blastp set to default parameters. For comparisons of amino acid sequences of greater than about 30 amino acids, the Blast 2 sequences function is employed using the default BLOSUM62 matrix set to default parameters, (gap existence cost of 11, and a per residue gap cost of 1). When aligning short peptides (fewer than around 30 amino acids), the alignment should be performed using the Blast 2 sequences function, employing the PAM30 matrix set to default parameters (open gap 9, extension gap 1 penalties). Proteins with even greater similarity to the reference sequence will show increasing percentage identities when assessed by this method, such as at least 70%, at least 75%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94% or at least 95% sequence identity. When less than the entire sequence is being compared for sequence identity, homologs will typically possess at least 75% sequence identity over short windows of 10-20 amino acids, and may possess sequence identities of at least 85% or at least 90% or 95% or more depending on their similarity to the reference sequence. Methods for determining sequence identity over such short windows are described at the NCBI web-site, frequently asked questions page. One of ordinary skill in the art will appreciate that these sequence identity ranges are provided for guidance only; it is entirely possible that strongly significant homologs could be obtained that fall outside of the ranges provided. The present invention provides not only the peptide homologs that are described above, but also nucleic acid molecules that encode such homologs. *ELF3* homologs will typically also have ELF3 protein biological activity.

An alternative indication that two nucleic acid molecules are closely related is that the two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence-dependent and are different under different environmental parameters. Generally, stringent conditions are selected to be about 5°C to 20°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Conditions for nucleic acid hybridization and calculation of stringencies can be found in Sambrook *et al.* (1989) and Tijssen (1993). Nucleic acid molecules that hybridize under stringent conditions to the *Arabidopsis ELF3* sequences will typically hybridize to a probe based on either the entire *Arabidopsis ELF3* cDNA or selected portions of the cDNA under wash conditions of 0.2 x SSC, 0.1% SDS at 55°C for 1 hour. A more detailed discussion of hybridization conditions, including low stringency conditions, is presented below.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences, due to the degeneracy of the genetic code. It is understood that changes in nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid molecules that all encode substantially the same protein.

Ortholog: Two nucleotide or amino acid sequences are orthologs of each other if they share a common ancestral sequence and diverged when a species carrying that ancestral sequence split into two species. Orthologous sequences are also homologous sequences.

Specific binding agent: An agent that binds substantially only to a defined target. Thus an ELF3 protein specific binding agent binds substantially only the ELF3 protein. As used herein, the term "ELF3 protein specific binding agent" includes anti- ELF3 protein antibodies and other agents that bind substantially only to the ELF3 protein.

Anti-ELF3 protein antibodies may be produced using standard procedures described in a number of texts, including Harlow and Lane (1988). The determination that a particular agent binds substantially only to the ELF3 protein may readily be made by using or adapting routine procedures. One suitable *in vitro* assay makes use of the Western blotting procedure (described in many standard texts, including Harlow and Lane (1988)). Western blotting may be used to determine that a given ELF3 protein binding agent, such as an anti-ELF3 protein monoclonal antibody, binds substantially only to the ELF3 protein.

Vector: A nucleic acid molecule as introduced into a host cell, thereby producing a transformed host cell. A vector may include nucleic acid sequences that permit it to replicate in a host cell, such as an origin of replication. A vector may also include one or more selectable marker genes and other genetic elements known in the art.

Transformed: A transformed cell is a cell into which has been introduced a nucleic acid molecule by molecular biology techniques. As used herein, the term transformation encompasses all techniques by which a nucleic acid molecule might be introduced into such a cell, including transfection with viral vectors, transformation with plasmid vectors, and introduction of naked DNA by electroporation, lipofection, and particle gun acceleration.

Isolated: An "isolated" biological component (such as a nucleic acid molecule, protein or organelle) has been substantially separated or purified away from other biological components in the cell of the organism in which the component naturally occurs, *i.e.*, other chromosomal and extra-chromosomal DNA and RNA, proteins and organelles. Nucleic acid molecules and proteins that have been "isolated" include nucleic acid molecules and proteins purified by standard purification methods. The term also embraces nucleic acid molecules and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acid molecules.

Purified: The term purified does not require absolute purity; rather, it is intended as a relative term. Thus, for example, a purified ELF3 protein preparation is one in which the ELF3 protein is more

enriched than the protein is in its natural environment within a cell. Generally, a preparation of ELF3 protein is purified such that ELF3 represents at least 5% of the total protein content of the preparation. For particular applications, higher purity may be desired, such that preparations in which ELF3 represents at least 25%, 50% or at least 90% of the total protein content may be employed.

Operably linked: A first nucleic acid sequence is operably linked with a second nucleic acid sequence when the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein-coding regions, in the same reading frame.

Recombinant: A recombinant nucleic acid is one that has a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two otherwise separated segments of sequence. This artificial combination can be accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, *e.g.*, by genetic engineering techniques.

cDNA (complementary DNA): A piece of DNA lacking internal, non-coding segments (introns) and regulatory sequences that determine transcription. cDNA is synthesized in the laboratory by reverse transcription from messenger RNA extracted from cells.

ORF (open reading frame): A series of nucleotide triplets (codons) coding for amino acids without any internal termination codons. These sequences are usually translatable into a peptide.

Transgenic plant: As used herein, this term refers to a plant that contains recombinant genetic material not normally found in plants of this type and which has been introduced into the plant in question (or into progenitors of the plant) by human manipulation. Thus, a plant that is grown from a plant cell into which recombinant DNA is introduced by transformation is a transgenic plant, as are all offspring of that plant that contain the introduced transgene (whether produced sexually or asexually).

II. ELF3 Protein and Nucleic Acid Sequences

This invention provides ELF3 proteins and *ELF3* nucleic acid molecules, including cDNA and gene sequences. The prototypical *ELF3* sequences are the *Arabidopsis* sequences, and the invention provides for the use of these sequences to produce transgenic plants, such as corn and rice plants, having increased or decreased levels of ELF3 protein.

a. Arabidopsis ELF3

The *Arabidopsis ELF3* genomic sequence is shown in SEQ ID NO: 3. The sequence comprises three introns and four exons, and encodes a protein that is 696 amino acids in length (SEQ ID NO: 2 shows the amino acid sequence of the ELF3 protein). The *Arabidopsis* ELF3 protein shares no

significant homology to any known published proteins with assigned function. However, one published *Arabidopsis* EST (GenBank # N96569; Newman *et al.*, 1994) overlaps nucleotides 853-2088 of the *Arabidopsis ELF3* open reading frame (ORF) (SEQ ID NO: 4) (nucleotides 1136-2501 of the *Arabidopsis ELF3* cDNA, SEQ ID NO: 1).

GenBank accession numbers for *ELF3* and putative *ELF3* homologs identified as such by this research group are as follows: *AtELF3* (*A. thaliana* genomic DNA: AC004747, published December 17, 1999), *AtEEC* (*A. thaliana* genomic DNA: AB023045, published November 20, 1999), *cELF3* (yet to be submitted), *tELF3* [*Lycopersicon esculentum* Expressed Sequence Tags (ESTs) from Clemson University Genomics Institute: AW093790 (October 18, 1999), AI894513 (July 27, 1999), AI488927 (June 29, 1999), AI486934 (June 29, 1999), AI894398 (July 27, 1999)], *rELF3* (*Oryza sativa* genomic DNA: AP000399, published December 3, 1999), and *mELF3* (*Zea mays* EST from Stanford University Genome Center: AI637184, published April 26, 1999).

The cDNA corresponding to the *ELF3* gene is shown in SEQ ID NO: 1, and the *ELF3* ORF is shown in SEQ ID NO: 4. As described below, the *Arabidopsis* ELF3 protein has ELF3 biological activity, *i.e.*, it complements the defective characteristics of photoperiod-insensitive flowering and elongated hypocotyl in *elf3* mutant plants when the *ELF3* gene sequence is introduced into these plants and the ELF3 protein is thereby expressed. In addition, ELF3 proteins contain one or more ESSENCE of ELF3 CONSENSUS (EEC) regions (see Figure 2).

With the provision herein of the *Arabidopsis ELF3* cDNA and gene sequences, the polymerase chain reaction (PCR) may now be utilized as a preferred method for producing nucleic acid sequences encoding the *Arabidopsis* ELF3 protein. For example, PCR amplification of the *Arabidopsis ELF3* cDNA sequence may be accomplished either by direct PCR from a plant cDNA library or by reverse-transcription PCR (RT-PCR) using RNA extracted from plant cells as a template. Methods and conditions for both direct PCR and RT-PCR are known in the art and are described in Innis *et al.* (1990). Any plant cDNA library would be useful for direct PCR. The *ELF3* gene sequences can be isolated from other libraries, for instance the IGF *Arabidopsis* BAC library (Mozo *et al.* 1998)

The selection of PCR primers will be made according to the portions of the *ELF3* cDNA (or gene) that are to be amplified. Primers may be chosen to amplify small segments of the cDNA, the open reading frame, the entire cDNA molecule or the entire gene sequence. Variations in amplification conditions may be required to accommodate primers of differing lengths; such considerations are well known in the art and are discussed in Innis *et al.* (1990), Sambrook *et al.* (1989), and Ausubel *et al.* (1992). By way of example only, the *Arabidopsis ELF3* cDNA molecule as shown in SEQ ID NO: 1 (excluding the poly A tail) may be amplified using the following combination of primers:

primer 1: 5' TGAAAACTCACTTTGGTTTTGTTTG 3' (SEQ ID NO: 6)
primer 2: 5' AAGACAAATTAACACATATAAATGA 3' (SEQ ID NO: 7)

The open reading frame portion of the cDNA (SEQ ID NO: 4) may be amplified using the following primer pair:

primer 3: 5' ATGAATAGAGGGAAAGATGAGGAG 3' (SEQ ID NO: 8)
primer 4: 5' TTAAGGCTTAGAGGAGTCATAGCGT 3' (SEQ ID NO: 9)

These primers are illustrative only; one of ordinary skill in the art will appreciate that many different primers may be derived from the provided cDNA and gene sequences in order to amplify particular regions of these molecules. Resequencing of PCR products obtained by these amplification procedures is recommended; this will facilitate confirmation of the amplified sequence and will also provide information on natural variation in this sequence in different ecotypes and plant populations. Oligonucleotides derived from the *Arabidopsis ELF3* sequence may be used in such sequencing methods.

Oligonucleotides that are derived from the *Arabidopsis ELF3* cDNA or gene sequences are encompassed within the scope of the present invention. Preferably, such oligonucleotide primers will comprise a sequence of at least 15-20 consecutive nucleotides of the *Arabidopsis ELF3* cDNA or gene sequences. To enhance amplification specificity, oligonucleotide primers comprising at least 25, 30, 35, 40, 45 or 50 consecutive nucleotides of these sequences may also be used.

b. *ELF3* Genes in Other Plant Species

Orthologs of the *ELF3* gene are present in a number of plant species including *Chlamydomonas*, Douglas fir, corn, broccoli, cauliflower, soybean, *Medicago*, rice, poplar, tobacco, *Cardamine*, and tomato (see Examples 4, 5 and 6, below). With the provision herein of the prototypical ELF3 protein from *Arabidopsis* and cDNA and gene sequences that encode this protein, cloning of cDNAs and genes that encode ELF3 protein orthologs in other plant species is now enabled. Standard methods, including those described herein, can be used. As described above, orthologs of the disclosed *Arabidopsis* ELF3 protein have ELF3 protein biological activity and typically possess at least 60% sequence identity counted over the full length alignment with the amino acid sequence of *Arabidopsis* ELF3 using the NCBI Blast 2.0, gapped blastp set to default parameters. Proteins with even greater similarity to the *Arabidopsis* sequence will show greater percentage identities when assessed by this method, such as at least 70%, at least 75%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94% or at least 95% or more sequence identity.

Both conventional hybridization and PCR amplification procedures may be utilized to clone sequences encoding ELF3 protein orthologs. Common to these techniques is the hybridization of

probes or primers derived from the *Arabidopsis ELF3* cDNA or gene sequence to a target nucleotide preparation. This target may be, in the case of conventional hybridization approaches, a cDNA or genomic library or, in the case of PCR amplification, a cDNA or genomic library, or an mRNA preparation.

Direct PCR amplification may be performed on cDNA or genomic libraries prepared from the plant species in question, or RT-PCR may be performed using mRNA extracted from the plant cells using standard methods. PCR primers will comprise at least 15 consecutive nucleotides of the *Arabidopsis ELF3* cDNA or gene. One of ordinary skill in the art will appreciate that sequence differences between the *Arabidopsis ELF3* cDNA or gene and the target nucleic acid to be amplified may result in lower amplification efficiencies. To compensate for this difference, longer PCR primers or lower annealing temperatures may be used during the amplification cycle. Where lower annealing temperatures are used, sequential rounds of amplification using nested primer pairs may be necessary to enhance amplification specificity.

For conventional hybridization techniques, the hybridization probe is preferably conjugated with a detectable label such as a radioactive label, and the probe is preferably of at least 20 nucleotides in length. As is well known in the art, increasing the length of hybridization probes tends to give enhanced specificity. The labeled probe derived from the *Arabidopsis* cDNA or gene sequence may be hybridized to a plant cDNA or genomic library and the hybridization signal detected using means known in the art. The hybridizing colony or plaque (depending on the type of library used) is then purified and the cloned sequence contained in that colony or plaque isolated and characterized.

Orthologs of the *Arabidopsis ELF3* may alternatively be obtained by immunoscreening an expression library. With the provision herein of the disclosed *Arabidopsis ELF3* nucleic acid sequences, the protein may be expressed in and purified from a heterologous expression system (*e.g.*, *E. coli*) and used to raise antibodies (monoclonal or polyclonal) specific for the *Arabidopsis* ELF3 protein. Antibodies may also be raised against synthetic peptides derived from the *Arabidopsis* ELF3 amino acid sequence presented herein. Methods of raising antibodies are well known in the art and are described in Harlow and Lane (1988). Such antibodies can be used to screen an expression cDNA library produced from the plant from which it is desired to clone the *ELF3* ortholog, using routine methods. The selected cDNAs can be confirmed by sequencing.

c. *ELF3* Sequence Variants

With the provision of the *Arabidopsis* ELF3 protein and *ELF3* cDNA and gene sequences herein, the creation of variants of these sequences is now enabled.

Variant ELF3 proteins include proteins that differ in amino acid sequence from the *Arabidopsis* ELF3 sequence disclosed but which retain ELF3 protein biological activity. Such proteins may be produced by manipulating the nucleotide sequence of the *Arabidopsis ELF3* cDNA or gene using

standard procedures, including for instance site-directed mutagenesis or PCR. The simplest modifications involve the substitution of one or more amino acids for amino acids having similar biochemical properties. These so-called conservative substitutions are likely to have minimal impact on the activity of the resultant protein. Table 1 shows amino acids that may be substituted for an original amino acid in a protein, and which are regarded as conservative substitutions.

Table 1.

Original Residue	Conservative Substitutions
Ala	ser
Arg	lys
Asn	gln; his
Asp	glu
Cys	ser
Gln	asn
Glu	asp
Gly	pro
His	asn; gln
Ile	leu; val
Leu	ile; val
Lys	arg; gln; glu
Met	leu; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

More substantial changes in protein functions or other features may be obtained by selecting amino acid substitutions that are less conservative than those listed in Table 1. Such changes include changing residues that differ more significantly in their effect on maintaining polypeptide backbone structure (*e.g.*, sheet or helical conformation) near the substitution, charge or hydrophobicity of the molecule at the target site, or bulk of a specific side chain. The following substitutions are generally expected to produce the greatest changes in protein properties: (a) a hydrophilic residue (*e.g.*, seryl or threonyl) is substituted for (or by) a hydrophobic residue (*e.g.*, leucyl, isoleucyl, phenylalanyl, valyl or alanyl); (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain (*e.g.*, lysyl, arginyl, or histadyl) is substituted for (or by) an electronegative residue (*e.g.*, glutamyl or aspartyl); or (d) a residue having a bulky side chain (*e.g.*, phenylalanine) is substituted for (or by) one lacking a side chain (*e.g.*, glycine). The effects of these amino acid substitutions, deletions, or additions may be assessed in ELF3 protein derivatives by analyzing the ability of a gene encoding the derivative protein to complement the photoperiod-insensitive flowering and elongated hypocotyl defects in an *elf3* mutant. Alternatively, the effect may be examined by

studying circadian influenced CAB-*luc* transcription and/or leaf movement as discussed in Example 2, below.

Variant *ELF3* cDNA or genes may be produced by standard DNA mutagenesis techniques, for example, M13 primer mutagenesis. Details of these techniques are provided in Sambrook *et al.* (1989), Ch. 15. By the use of such techniques, variants may be created which differ in minor ways from the *Arabidopsis ELF3* cDNA or gene sequences disclosed, yet which still encode a protein having ELF3 protein biological activity. DNA molecules and nucleotide sequences that are derivatives of those specifically disclosed herein and that differ from those disclosed by the deletion, addition, or substitution of nucleotides while still encoding a protein that has ELF3 protein biological activity are comprehended by this invention. In their most simple form, such variants may differ from the disclosed sequences by alteration of the coding region to fit the codon usage bias of the particular organism into which the molecule is to be introduced.

Alternatively, the coding region may be altered by taking advantage of the degeneracy of the genetic code to alter the coding sequence such that, while the nucleotide sequence is substantially altered, it nevertheless encodes a protein having an amino acid sequence substantially similar to the disclosed *Arabidopsis* ELF3 protein sequence. For example, the 23rd amino acid residue of the *Arabidopsis* ELF3 protein is alanine. This alanine residue is encoded for by the nucleotide codon triplet GCA. Because of the degeneracy of the genetic code, three other nucleotide codon triplets - GCT, GCC and GCG - also code for alanine. Thus, the nucleotide sequence of the *Arabidopsis ELF3* ORF could be changed at this position to any of these three alternative codons without affecting the amino acid composition or other characteristics of the encoded protein. Based upon the degeneracy of the genetic code, variant DNA molecules may be derived from the cDNA and gene sequences disclosed herein using standard DNA mutagenesis techniques as described above, or by synthesis of DNA sequences. Thus, this invention also encompasses nucleic acid sequences that encode an ELF3 protein, but which vary from the disclosed nucleic acid sequences by virtue of the degeneracy of the genetic code.

Variants of the ELF3 protein may also be defined in terms of their sequence identity with the prototype ELF3 protein shown in SEQ ID NO: 2. As described above, ELF3 proteins have ELF3 biological activity and share at least 60% sequence identity with the *Arabidopsis* ELF3 protein. Nucleic acid sequences that encode such proteins may readily be determined simply by applying the genetic code to the amino acid sequence of an ELF3 protein, and such nucleic acid molecules may readily be produced by assembling oligonucleotides corresponding to portions of the sequence.

Nucleic acid molecules that are derived from the *Arabidopsis ELF3* cDNA and gene sequences disclosed include molecules that hybridize under stringent conditions to the disclosed prototypical *ELF3* nucleic acid molecules, or fragments thereof. Stringent conditions are hybridization at 55°C in 6 x SSC, 5 x Denhardt's solution, 0.1% SDS and 100 μg sheared salmon testes DNA, followed by 15-30 minute

sequential washes at 55°C in 2 x SSC, 0.1% SDS, followed by 1 x SSC, 0.1% SDS and finally 0.2 x SSC, 0.1% SDS.

Low stringency hybridization conditions (to detect less closely related homologs) are performed as described above but at 50°C (both hybridization and wash conditions); however, depending on the strength of the detected signal, the wash steps may be terminated after the first 2 x SSC, 0.1% SDS wash.

The *Arabidopsis ELF3* gene or cDNA, and orthologs of these sequences from other plants, may be incorporated into transformation vectors and introduced into plants to produce plants with an altered photoperiodic or circadian rhythm phenotype, as described below.

III. Introducing *ELF3* into Plants

Once a nucleic acid molecule (*e.g.*, cDNA or gene) encoding a protein involved in the determination of a particular plant characteristic has been isolated, standard techniques may be used to express the cDNA in transgenic plants in order to modify that particular plant characteristic. The basic approach is to clone, for instance, the cDNA into a transformation vector, such that it is operably linked to control sequences (*e.g.*, a promoter) that direct expression of the cDNA in plant cells. The transformation vector is then introduced into plant cells by one of a number of techniques (*e.g.*, electroporation) and progeny plants containing the introduced cDNA are selected. Preferably all or part of the transformation vector will stably integrate into the genome of the plant cell. That part of the transformation vector that integrates into the plant cell and that contains the introduced cDNA and associated sequences for controlling expression (the introduced "transgene") may be referred to as the recombinant expression cassette.

Selection of progeny plants containing the introduced transgene may be based upon the detection of an altered phenotype. Such a phenotype may result directly from the cDNA cloned into the transformation vector or may be manifested as enhanced resistance to a chemical agent (such as an antibiotic) as a result of the inclusion of a dominant selectable marker gene incorporated into the transformation vector.

Successful examples of the modification of plant characteristics by transformation with cloned cDNA sequences are replete in the technical and scientific literature. Selected examples, which serve to illustrate the knowledge in this field of technology, include:

U.S. Patent No. 5,451,514 (modification of lignin synthesis using antisense RNA and co-suppression);

U.S. Patent No. 5,750,385 (modification of plant light-, seed- and fruit-specific gene expression using sense and antisense transformation constructs);

U.S. Patent No. 5,583,021 (modification of virus resistance by expression of plus-sense untranslatable RNA);

U.S. Patent No. 5,589,615 (production of transgenic plants with increased nutritional value via the expression of modified 2S storage albumins);

U.S. Patent No. 5,268,526 (modification of phytochrome expression in transgenic plants);

U.S. Patent No. 5,741,684 (production of plants resistant to herbicides or antibiotics through the use of anti-sense expression);

U.S. Patent No. 5,773,692 (modification of the levels of chlorophyll by transformation of plants with anti-sense messages corresponding to *chlorophyll a/b* binding protein);

WO 96/13582 (modification of seed VLCFA composition using over expression, co-suppression and antisense RNA in conjunction with the Arabidopsis *FAE1* gene)

These examples include descriptions of transformation vector selection, transformation techniques and the assembly of constructs designed to over-express the introduced nucleic acid, as well as techniques for sense suppression and antisense expression. In light of the foregoing and the provision herein of the *Arabidopsis ELF3* cDNA and gene sequences, one of ordinary skill in the art will be able to introduce these nucleic acid molecules, or orthologous, homologous or derivative forms of these molecules, into plants in order to produce plants having altered ELF3 activity. Manipulating the expression of ELF3 in plants will be useful to confer altered circadian clock and/or photoperiodism function. Alteration of the ELF3 protein levels in plants could be used to re-set or customize the circadian clock, for instance in order to alter the plant developmental patterns or photoperiodic responses (*e.g.*, the timing of floral development).

a. Plant Types

The presence of a circadian cycle appears to be universal, occurring not only in all plants thus far examined, but also in insects, including *Drosophila* (Hall, 1990) and microbes such as *Neurospora crassa* (Dunlap, 1993). At the molecular level, *ELF3* homologs have been found in a variety of plant species (see Example 4, below). Thus, expression of the ELF3 protein may be modified in a wide range of higher plants to confer altered circadian clock and/or photoperiodism function, including monocotyledonous and dicotyledenous plants. These include, but are not limited to, *Arabidopsis*, *Cardamine*, cotton, tobacco, maize, wheat, rice, barley, soybean, beans in general, rape/canola, alfalfa, flax, sunflower, safflower, brassica, cotton, flax, peanut, clover; vegetables such as lettuce, tomato, cucurbits, cassava, potato, carrot, radish, pea, lentils, cabbage, cauliflower, broccoli, Brussels sprouts, peppers; tree fruits such as citrus, apples, pears, peaches, apricots, walnuts; other trees including poplar, oak, maple, pine, spruce and other conifers; and flowers or other ornamental plants such as carnations, roses, petunias, orchids, impatiens, pansies, lilies, snapdragons, geraniums, and so forth.

b. Vector Construction, Choice of Promoters

A number of recombinant vectors suitable for stable transformation of plant cells or for the establishment of transgenic plants have been described including those described in Pouwels *et al.*, (1987), Weissbach and Weissbach, (1989), and Gelvin *et al.*, (1990). Typically, plant transformation vectors include one or more cloned plant genes (or cDNAs) under the transcriptional control of 5' and 3' regulatory sequences; and at least one dominant selectable marker. Such plant transformation vectors typically also contain a promoter regulatory region (*e.g.*, a regulatory region controlling inducible or constitutive, environmentally-or developmentally-regulated, or cell- or tissue-specific expression), a transcription initiation start site, a ribosome binding site, an RNA processing signal, a transcription termination site, and/or a polyadenylation signal.

Examples of constitutive plant promoters that may be useful for expressing an *ELF3* nucleic acid molecule include: the cauliflower mosaic virus (CaMV) 35S promoter, which confers constitutive, high-level expression in most plant tissues (*see, e.g.*, Odel *et al.*, 1985, Dekeyser *et al.*, 1990, Terada and Shimamoto, 1990; Benfey and Chua. 1990); the nopaline synthase promoter (An *et al.*, 1988); and the octopine synthase promoter (Fromm *et al.*, 1989).

A variety of plant gene promoters are regulated in response to environmental, hormonal, chemical, and/or developmental signals, and can be used for expression of the cDNA in plant cells. Such promoters include, for instance, those regulated by: (a) heat (Callis *et al.*, 1988; Ainley, *et al.* 1993; Gilmartin *et al.* 1992); (b) light (*e.g.*, the pea rbcS-3A promoter, Kuhlemeier *et al.*, 1989, and the maize rbcS promoter, Schaffner and Sheen, 1991); (c) hormones, such as abscisic acid (Marcotte *et al.*, 1989); (d) wounding (*e.g.*, *wunI*, Siebertz *et al.*, 1989); and (e) chemicals such as methyl jasminate or salicylic acid (*see also* Gatz *et al.*, 1997).

Alternatively, tissue specific (root, leaf, flower, or seed, for example) promoters (Carpenter *et al.* 1992, Denis *et al.* 1993, Opperman *et al.* 1993, Stockhause *et al.* 1997; Roshal *et al.*, 1987; Schernthaner *et al.*, 1988; and Bustos *et al.*, 1989) can be fused to the coding sequence to obtained protein expression in specific organs.

Promoters responsive to the circadian cycle can also be used in plant gene expression vectors. Such promoters include the native *ELF3* promoter as described herein, and the promoter from the *chlorophyll a/b* binding protein (Millar *et al.* 1992).

Plant transformation vectors may also include RNA processing signals, for example, introns, which may be positioned upstream or downstream of the ORF sequence in the transgene. In addition, the expression vectors may include further regulatory sequences from the 3'-untranslated region of plant genes, *e.g.*, a 3' terminator region to increase mRNA stability of the mRNA, such as the PI-II terminator region of potato or the *Agrobacterium* octopine or nopaline synthase 3' terminator regions. The 3' region of the *ELF3* gene can also be used.

Finally, as noted above, plant transformation vectors may include dominant selectable marker genes to allow for the ready selection of transformants. Such genes include those encoding antibiotic resistance genes (*e.g.*, resistance to hygromycin, kanamycin, bleomycin, G418, streptomycin or spectinomycin) and herbicide resistance genes (*e.g.*, phosphinothricin acetyltransferase).

c. Arrangement of *ELF3* Sequence in the Vector

The particular arrangement of the *ELF3* sequence in the transformation vector will be selected according to the type of expression of the sequence that is desired.

Where enhanced ELF3 protein activity is desired in the plant, an *ELF3* ORF may be operably linked to a constitutive high-level promoter such as the CaMV 35S promoter. As noted below, modification of *ELF3* synthesis may also be achieved by introducing into a plant a transformation vector containing a variant form of an *ELF3* cDNA or gene.

In contrast, a reduction of ELF3 activity in the transgenic plant may be obtained by introducing into plants an antisense construct based on an *ELF3* cDNA or gene sequence. For antisense suppression, an *ELF3* cDNA or gene is arranged in reverse orientation relative to the promoter sequence in the transformation vector. The introduced sequence need not be a full length *ELF3* cDNA or gene, and need not be exactly homologous to the native *ELF3* cDNA or gene found in the plant type to be transformed. Generally, however, where the introduced sequence is of shorter length, a higher degree of homology to the native *ELF3* sequence will be needed for effective antisense suppression. The introduced antisense sequence in the vector generally will be at least 30 nucleotides in length, and improved antisense suppression will typically be observed as the length of the antisense sequence increases. Preferably, the length of the antisense sequence in the vector will be greater than 100 nucleotides. Transcription of an antisense construct as described results in the production of RNA molecules that are the reverse complement of mRNA molecules transcribed from the endogenous *ELF3* gene in the plant cell. Although the exact mechanism by which antisense RNA molecules interfere with gene expression has not been elucidated, it is believed that antisense RNA molecules bind to the endogenous mRNA molecules and thereby inhibit translation of the endogenous mRNA. The production and use of anti-sense constructs are disclosed, for instance, in U.S. Pat. Nos. 5,773,692 (using constructs encoding anti-sense RNA for *chlorophyll a/b* binding protein to reduce plant chlorophyll content), and 5,741,684 (regulating the fertility of pollen in various plants through the use of anti-sense RNA to genes involved in pollen development or function).

Suppression of endogenous *ELF3* gene expression can also be achieved using ribozymes. Ribozymes are synthetic RNA molecules that possess highly specific endoribonuclease activity. The production and use of ribozymes are disclosed in U.S. Patent No. 4,987,071 to Cech and U.S. Patent No. 5,543,508 to Haselhoff. Inclusion of ribozyme sequences within antisense RNAs may be used to confer RNA cleaving activity on the antisense RNA, such that endogenous mRNA molecules that bind

to the antisense RNA are cleaved, leading to an enhanced antisense inhibition of endogenous gene expression.

Constructs in which an *ELF3* cDNA or gene (or variants thereof) are over-expressed may also be used to obtain co-suppression of the endogenous *ELF3* gene in the manner described in U.S. Patent No. 5,231,021 to Jorgensen. Such co-suppression (also termed sense suppression) does not require that the entire *ELF3* cDNA or gene be introduced into the plant cells, nor does it require that the introduced sequence be exactly identical to the endogenous *ELF3* gene. However, as with antisense suppression, the suppressive efficiency will be enhanced as (1) the introduced sequence is lengthened and (2) the sequence similarity between the introduced sequence and the endogenous *ELF3* gene is increased.

Constructs expressing an untranslatable form of an *ELF3* mRNA may also be used to suppress the expression of endogenous ELF3 activity. Methods for producing such constructs are described in U.S. Patent No. 5,583,021 to Dougherty *et al.* Preferably, such constructs are made by introducing a premature stop codon into an *ELF3* ORF.

Finally, dominant negative mutant forms of the disclosed sequences may be used to block endogenous *ELF3* activity. Such mutants require the production of mutated forms of the ELF3 protein that interact with the same molecules as ELF3 but do not have ELF3 activity.

d. Transformation and Regeneration Techniques

Transformation and regeneration of both monocotyledonous and dicotyledonous plant cells is now routine, and the most appropriate transformation technique will be determined by the practitioner. The choice of method will vary with the type of plant to be transformed; those skilled in the art will recognize the suitability of particular methods for given plant types. Suitable methods may include, but are not limited to: electroporation of plant protoplasts; liposome-mediated transformation; polyethylene glycol (PEG) mediated transformation; transformation using viruses; micro-injection of plant cells; micro-projectile bombardment of plant cells; vacuum infiltration; and *Agrobacterium tumefaciens* (AT) mediated transformation. Typical procedures for transforming and regenerating plants are described in the patent documents listed at the beginning of this section.

e. Selection of Transformed Plants

Following transformation and regeneration of plants with the transformation vector, transformed plants are usually selected using a dominant selectable marker incorporated into the transformation vector. Typically, such a marker will confer antibiotic resistance on the seedlings of transformed plants, and selection of transformants can be accomplished by exposing the seedlings to appropriate concentrations of the antibiotic.

After transformed plants are selected and grown to maturity, they can be assayed using the methods described herein to determine whether the circadian cycle or photoperiodism of the transformed plant has been altered as a result of the introduced transgene.

IV. Production of Recombinant ELF3 Protein in Heterologous Expression Systems

Many different expression systems are available for expressing cloned nucleic acid molecules. Examples of prokaryotic and eukaryotic expression systems that are routinely used in laboratories are described in Chapters 16-17 of Sambrook *et al.* (1989). Such systems may be used to express ELF3 at high levels to facilitate purification of the protein. The purified ELF3 protein may be used for a variety of purposes. For example, the purified recombinant enzyme may be used as an immunogen to raise anti- ELF3 antibodies. Such antibodies are useful as both research reagents (such as in the study of circadian clock and photoperiodism mechanisms in plants) as well as diagnostically to determine expression levels of the protein in plants that are being developed for agricultural or other use. Thus, the antibodies may be used to quantify the level of ELF3 protein both in non-transgenic plant varieties and in transgenic varieties that are designed to over-express or under-express the ELF3 protein. Such quantification may be performed using standard immunoassay techniques, such as ELISA and *in situ* immunofluorescence and others described in Harlow & Lane (1988).

By way of example only, high level expression of the ELF3 protein may be achieved by cloning and expressing the *ELF3* cDNA in yeast cells using the pYES2 yeast expression vector (INVITROGEN, Carlsbad, CA). Alternatively, a genetic construct may be produced to direct secretion of the recombinant ELF3 protein from yeast cells into the growth medium. This approach will facilitate the purification of the ELF3 protein, if this is necessary. Secretion of the recombinant protein from the yeast cells may be achieved by placing a yeast signal sequence adjacent to the *ELF3* coding region. A number of yeast signal sequences have been characterized, including the signal sequence for yeast invertase. This sequence has been successfully used to direct the secretion of heterologous proteins from yeast cells, including such proteins as human interferon (Chang *et al.*, 1986), human lactoferrin (Liang and Richardson, 1993) and prochymosin (Smith *et al.*, 1985).

Alternatively, the enzyme may be expressed at high level in prokaryotic expression systems, such as *E. coli*, as described in Sambrook *et al.* (1989). Commercially available prokaryotic expression systems include the pBAD expression system and the ThioFusion expression system (INVITROGEN, Carlsbad, CA).

V. ELF3 Promoter

The 5' regulatory region of the *ELF3* gene is also provided herein (SEQ ID NO: 5). This regulatory region confers *ELF3*-like circadian rhythm-based expression on open reading frames to which it is operably linked. Approximately 4 kb of the *ELF3* 5' regulatory region is provided in SEQ ID NO: 5. While this entire *ca.* 4 kb regulatory sequence may be employed, one of ordinary skill in the

art will appreciate that less than this entire sequence may be sufficient to confer *ELF3*-like circadian rhythm expression. For example, sequences comprising nucleotides 1-4071 of SEQ ID NO: 5 or shorter sequences such as those spanning nucleotides 500-4071, 1000-4071, 1500-4071, 2000-4071, 2500-4071, 3000-4071, 3500-4071 and 4000-4071 may be employed. One embodiment of an ELF3 promoter is about nucleotides 1 through about 1900 of the 5' upstream region shown in SEQ ID NO: 5. Other particular embodiments include about nucleotides 50-1900, 150-1900, 250-1900, 350-1900, 450-1900, 550-1900 and so forth.

Sequences as short as 50 or 100 nucleotides from within the 5' regulatory region of ELF3 may also be employed. The determination of whether a particular sub-region of the disclosed sequence operates to confer effective ELF3-like circadian rhythm expression in a particular system (taking into account the plant species into which the construct is being introduced, the level of expression required, etc.) will be performed using known methods. These include, for instance, operably linking the promoter sub-region to a marker gene (*e.g.* GUS or luciferase), introducing such constructs into plants, and determining the level of expression of the marker gene.

The present invention therefore facilitates the production, by standard molecular biology techniques, of nucleic acid molecules comprising this promoter sequence operably linked to a nucleic acid sequence, such as an open reading frame. Suitable open reading frames include open reading frames encoding any protein for which *ELF3*-like circadian rhythm expression is desired.

EXAMPLES

Example 1: Cloning *Arabidopsis ELF3*

The *ELF3* gene was isolated by map-based positional cloning. Molecular markers tightly linked to the *ELF3* gene were identified by random fragment length polymorphism (RFLP) analysis, and a high resolution genetic map of the locus was constructed. The region containing the *ELF3* gene was narrowed down to 30 kb contained on a single bacterial artificial chromosome (BAC). This BAC was sequenced, and cDNAs with homology to sequences within the BAC were isolated from a variety of cDNA libraries. The ELF3 sequence was further localized by complementation experiments to a 10 kb subcloned fragment contained within the BAC. Identification of the appropriate gene within the subcloned fragment was confirmed through isolation and sequencing of *elf3* alleles from various *Arabidopsis elf3* mutants.

The isolated *ELF3* gene (SEQ ID NO: 3) has no significant sequence similarity to other DNA or protein sequences with assigned function. However, a published EST (GenBank # N96569; Newman *et al.*, 1994) overlaps nucleotide 1235-2501 of the corresponding cDNA (SEQ ID NO: 1).

ELF3 has four exons, and is transcribed as an mRNA of about 2.4 kb in *Arabidopsis* seedlings and in mature leaves. The putative protein (SEQ ID NO: 2) encoded by the *ELF3* ORF (SEQ ID NO: 4) is 695 amino acids in length and has a predicted molecular weight of approximately 80 KDa.

Research by this group has recently identified several putative ELF3 orthologs from other plant species, including *Cardamine oligosperma*, tomato, rice, and maize (see Examples 4 and 5, below). GenBank accession numbers for *ELF3* and putative *ELF3* homologs identified as such by this research group are as follows: *AtELF3* (*A. thaliana* genomic DNA: AC004747, published December 17, 1999), *AtEEC* (*A. thaliana* genomic DNA: AB023045, published November 20, 1999), *cELF3* (yet to be submitted), *tELF3* [*Lycopersicon esculentum* Expressed Sequence Tags (ESTs) from Clemson University Genomics Institute: AW093790 (October 18, 1999), AI894513 (July 27, 1999), AI488927 (June 29, 1999), AI486934 (June 29, 1999), AI894398 (July 27, 1999)], *rELF3* (*Oryza sativa* genomic DNA: AP000399, published December 3, 1999), and *mELF3* (*Zea mays* EST from Stanford University Genome Center: AI637184, published April 26, 1999).

Example 2: Analysis of ELF3 Phenotype

Sensitive assays for monitoring circadian rhythm responses in *Arabidopsis* have been developed (Millar and Kay, 1991; Millar *et al.*, 1992). One assay system is based on the observation that the transcription of the *chlorophyll a/b* binding protein gene, *CAB2*, cycles on a 24-hour period. Transcription from the *CAB2* promoter increases prior to subjective dawn, peaks in late morning, and falls to a low level late in the day (Millar and Kay, 1991). Cycling of *CAB* mRNA continues under constant light conditions. In order to follow expression in vivo, the *CAB2* promoter has been fused to the gene encoding firefly luciferase (*luc*), and this fusion has been transformed in wild-type *Arabidopsis* (Millar *et al.*, 1992). Transcriptional expression from the *CAB2-luc* fusion construct is monitored by imaging single transgenic seedlings using a low-light video camera and a photon-counting image processor; the results from imaging the *CAB2-luc* fusion is comparable to the transcriptional expression of the endogenous *CAB2* gene. With this system, over one hundred individual seedlings can be imaged every 30 minutes, thus allowing the collection of thousands of data points in less than one week. This very powerful system has recently been used to characterize several known photomorphogenic *Arabidopsis* mutants (Millar *et al.*, 1995a) and to isolate a short-period mutant of *Arabidopsis* (Millar *et al.*, 1995b). *Elf3* mutants examined using this system are defective in circadian regulated *CAB2* transcription (Hicks *et al.*, 1996).

An automatic video imaging system can also be used to monitor a second circadian regulated process, leaf movement (Millar and Kay, 1991). Plant leaves turn down (open) during the day and turn up (closed) during the night in a circadian fashion. *Arabidopsis* seedlings display this circadian leaf movement in constant light, and this can be assayed and quantified using a relatively inexpensive video

and computer system (Millar and Kay, 1991). The analysis of leaf movements provides an independent circadian regulated process with which to evaluate potential circadian rhythm mutants (see, for instance, Schaffer *et al.* 1998, using leaf movement to analyze circadian cycle disruption in *late elongated hypocotyl* (*lhy*) mutants in *Arabidopsis*). *Elf3* mutants are also defective in circadian regulated leaf movements.

These assays may be used to assess the effect that modifying ELF3 protein expression level (*e.g.*, through introduction of *ELF3* antisense or sense constructs into plants) has on plant phenotype.

Example 3: Introducing *ELF3* Sequences into Plants

Plasmid construction

Arabidopsis ELF3 cDNA (SEQ ID NO: 1) and full-length genomic (SEQ ID NO: 3) sequences were used in the construction of over-expression and antisense vectors. These sequences were operably linked to the CaMV 35S (constitutive) promoter, in both the sense and antisense orientations, and cloned using standard molecular biology techniques into pSJL4 (Jones *et al.* 1992).

The over-expression and antisense expression cassettes were removed from the above vectors and inserted into pMON505 for *Agrobacterium*-mediated plant transformation.

Plant Transformation

Wild-type and *elf3* mutant *Arabidopsis* plants (ecotype Columbia) were transformed using standard *in planta Agrobacterium*-mediated techniques (Chang *et al.* 1994, Katavic *et al.* 1994). Transformed seeds were selected on kanamycin, and Kan^R seedlings transferred to soil and grown for further analysis.

Over-expression of ELF3 protein in *elf3* mutant plants comprising the ELF3 genomic gene sequence as the transgene resulted in full complementation of the *elf3* mutant phenotype in some transformed plants. In some instances, over-expression of ELF3 protein from cDNA-based transgenes in wild-type plants produced *elf3* mutant-like plants or plants having intermediate phenotype; this is probably the result of co-suppression. Antisense expression of the full-length *ELF3* cDNA in wild-type plants produced some transformants with an *elf3* mutant-like phenotype.

Example 4: *ELF3* Orthologs

As noted above, orthologs of *ELF3* exist in a number of plant species including corn, tomato and tobacco. The existence of these sequences may be demonstrated by hybridization techniques, such as Southern blotting. Hybridization was performed using a probe based on the entire *ELF3* cDNA sequence (SEQ ID NO: 1). This probe was hybridized to genomic DNA from *Arabidopsis*, *Chlamydomonas*, Douglas fir, corn, rice, poplar, tobacco, and tomato. High stringency hybridization was performed at 55°C in 6 x SSC, 5 x Denhardt's solution, 0.1% SDS and 100 µg sheared salmon

testes DNA, followed by 15-30 minute sequential washes at 55°C in 2 x SSC, 0.1% SDS, followed by 1 x SSC, 0.1% SDS and finally 0.2 x SSC, 0.1% SDS. A single, clean hybridizing band was observed on the Southern blot in *Arabidopsis*, rice, and tobacco genomic DNA preparations.

Lower stringency hybridization conditions were used to detect less closely related *ELF3* homologs. Such hybridization was performed at 50° C for 24 hours in the hybridization solution described above, followed by washing in 2 x SSC, 0.1% SDS at 50° C for 3 hours, with five sequential changes of wash solution. Hybridization of full length cDNA probe under low stringency hybridization conditions detected *ELF3* homologs (indicated by one or more bands on the Southern) in *Arabidopsis*, *Chlamydomonas*, Douglas fir, corn, rice, poplar, tobacco, and tomato and other plant species.

Once an *ELF3*-hybridizing band is detected in a plant species, standard techniques such as screening cDNA or genomic libraries from the plant with the *ELF3* probe may be used. Alternatively, *ELF3* homologs may be isolated by screening an expression library from the plant in question using a ELF3 protein specific binding agent, such as an anti-ELF3 antibody produced as described above. Such homologs may be introduced into plants using the methods described above in order to produce altered circadian rhythm and/or photoperiodic phenotypes.

It is also possible to use primers complementary to the *Arabidopsis ELF3* sequence to amplify orthologous nucleic acid sequences. For example, an *ELF3* ortholog has been isolated in this manner from a *Cardamine* genomic DNA preparation, using the following PCR amplification primers:

primer 5: 5' ATGAAGAGAGGGAAAGATGAGG 3' (SEQ ID NO:10)

primer 6: 5' GCCACCATCTCGGTATAACC 3' (SEQ ID NO:11).

Degenerate mixtures of oligonucleotides may also be used to amplify orthologous nucleic acid sequences. The construction of degenerate oligonucleotides is well known to one of ordinary skill in the art.

Nucleotide sequences from *C. oligosperma* (a member of the family Brassicaceae) were obtained by sequencing polymerase chain reaction products using degenerate oligonucleotides to the *Arabidopsis ELF3* gene and genomic DNA or cDNA prepared from *C. oligosperma* seedlings using standard techniques. The sequence of the amplified *Cardamine* ELF3 ortholog (cELF3) is shown in SEQ ID NO: 12.

Example 5: Consensus Sequences Within the ELF3 Protein and Homologs Thereof

Computerized, searchable databases were searched for sequences having significant homology the *Arabidopsis* ELF3 cDNA and genomic nucleotide sequences depicted herein, and the *Cardamine* ELF3 ortholog nucleotide sequence (SEQ ID NO: 12).

This search yielded several putative ELF3 homologs. GenBank accession numbers for *ELF3* and the putative *ELF3* homologs identified as such by this research group are as follows: *AtELF3* (*A. thaliana* genomic DNA: AC004747), *AtEEC* (*A. thaliana* genomic DNA: AB023045), *cELF3* (yet to be submitted), *tELF3* (*Lycopersicon esculentum* Expressed Sequence Tags (ESTs) from Clemson University Genomics Institute: AW093790, AI894513, AI488927, AI486934, AI894398), *rELF3* (*Oryza sativa* genomic DNA: AP000399), and *mELF3* (*Zea mays* EST from Stanford University Genome Center: AI637184).

Multiple sequence alignment of the ELF3 proteins (Figures 1 and 2) shows four highly conserved regions within ELF3 and putative ELF3 homologs from *Arabidopsis thaliana* (*Essence of ELF3 Consensus, EEC*) and other plant species (*Cardamine oligosperma*, tomato, rice, and maize) (Figure 2). Sequences were aligned and analyzed using CLUSTAL W (Thompson *et al.*, *Nucleic Acids Res.* 22, 4673-80, 1994) and PrettyBox (Genetics Computer Group, Inc.). Protein designations are given on the left. Amino acid residues are numbered on both the left and right. Residues shaded in black indicate identity of at least three ELF3/ELF3-related sequences in the alignment; light-shaded residues indicate similarity to consensus.

Example 6: Additional *ELF3* Orthologs

The ELF3 sequences and consensus sequences isolated as described above were used additional similar sequences from other plant species, using nucleic acid amplification and/or computer database searches. Additional ELF3 orthologs have been identified in *Medicago trunculata* (SEQ ID NOs: 36-39), *Pisum sativa* (SEQ ID NOs: 40-43), *Glycine max* (SEQ ID NOs: 44- and 45), *Xanthium* (SEQ ID NOs: 46-51), Poplar (SEQ ID NOs: 52-53), *Mimulus* (SEQ ID NOs: 54 and 55), *Zea mays* (SEQ ID NOs: 56 and 57), *Lycopersicon esculentum* (SEQ ID NOs: 58 and 59), and Broccoli (SEQ ID NOs: 60-63). Nucleic acid amplification, particularly polymerase chain amplification (PCR) also was used to confirm several of these sequences. For isolation and/or confirmation, amplification reactions were annealed at 55°C and extended for 35 seconds per round. The primers used were as follows:

Amplified ortholog	Forward primer	Reverse primer
Pisum sativa (SEQ ID NO: 40)	B-FWD, SEQ ID NO: 66	C-REV, SEQ ID NO: 68
Pisum sativa (SEQ ID NO: 42) first round	Pea1b-C-FWD, SEQ ID NO: 67	D-REV, SEQ ID NO: 65
Pisum sativa (SEQ ID NO: 42) second round	C-FWD, SEQ ID NO: 64	D-REV, SEQ ID NO: 65
Xanthium (SEQ ID NO: 46)	B-FWD, SEQ ID NO: 66	C-REV, SEQ ID NO: 68
Xanthium (SEQ ID NO: 48)	B-FWD, SEQ ID NO: 66	C-REV, SEQ ID NO: 68
Xanthium (SEQ ID NO: 50)	C-FWD, SEQ ID NO: 64	D-REV, SEQ ID NO: 65
Poplar (SEQ ID NO: 52)	C-FWD, SEQ ID NO: 64	D-REV, SEQ ID NO: 65

Mimulus (SEQ ID NO: 54)	C-FWD, SEQ ID NO: 64	D-REV, SEQ ID NO: 65
Zea mays (SEQ ID NO: 56)	B-FWD, SEQ ID NO: 66	C-REV, SEQ ID NO: 68
Zea mays (SEQ ID NO: 56)	C-FWD, SEQ ID NO: 64	D-REV, SEQ ID NO: 65
Broccoli (SEQ ID NO: 62)	C-FWD, SEQ ID NO: 64	D-REV, SEQ ID NO: 65

The amplified products were of the expected sizes.

The foregoing examples are provided by way of illustration only. One of skill in the art will appreciate that numerous variations on the biological molecules and methods described above may be employed to make and use the *ELF3* gene, corresponding protein, and promoter region. We claim all such subject matter that falls within the scope and spirit of the following claims.

References

Ainley *et al.* (1993) *Plant Mol. Biol.* 22:13-23.
Altschul *et al.* (1990). *J. Mol. Biol.*, 215, 403-10
Altschul *et al.* (1994). *Nature Genet.*, 6, 119-29.
An *et al.* (1988) *Plant Physiol.* 88:547.
Aronson *et al.* (1994) *Science* 263:1578-1584.
Ausubel *et al.* (1987) In *Current Protocols in Molecular Biology*, Greene Publishing Associates and Wiley-Intersciences.
Benfey and Chua (1990) *Science* 250:959-966.
Bernier (1988) *Ann. Rev. Plant Phys. and Plant Mol. Bio.* 39:175-219.
Bustos *et al.* (1989) *Plant Cell* 1:839.
Callis *et al.* (1988) *Plant Physiol.* 88:965.
Carpenter *et al.* (1992) *The Plant Cell* 4:557-571.
Chang *et al.* (1994) *Plant J.* 5:551-558.
Chang *et al.* (1986) *Mol. And Cell. Biol.* 6:1812-1819.
Corpet *et al.* (1988). *Nucleic Acids Research* 16, 10881-90.
Dekeyser *et al.* (1990) *Plant Cell* 2:591.
Denis *et al.* (1993) *Plant Physiol.* 101:1295-1304.
Dunlap (1993) *Annu. Rev. Physiol.* 55:683.
Edery *et al.* (1994) *Science* 263:237-240.
Fromm *et al.* (1989) *Plant Cell* 1:977.
Gatz *et al.* (1997) *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 48:89-108.
Gelvin *et al.* (1990) *Plant Molecular Biology Manual*, Kluwer Academic Publishers.
Gilmartin *et al.* (1992) *The Plant Cell* 4:839-949.
Hall (1990) *Ann. Rev. Genet.* 24:659.
Harlow & Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbor Laboratory, New York.
Hicks *et al.* (1996) *Science* 274(5288):790-792.
Higgins and Sharp (1988). *Gene*, 73: 237-244.
Higgins and Sharp (1989). *CABIOS* 5: 151-153.
Huang *et al.* (1992). *Computer Applications in the Biosciences* 8, 155-65.
Hülskamp *et al.* (1990) *Nature* 346:577-580.
Innis *et al.* (eds.) (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc., San Diego, California.
Jones *et al.* (1992) *Transgenic Res.* 1:285-297.
Katavic *et al.* (1994) *Mol. Gen. Genet.* 245:363-370.
Koornneef *et al.* (1991) *Mol. Gen. Genet.* 229:57-66.
Kuhlemeier *et al.* (1989) *Plant Cell* 1: 471.
Lambie and Kimble (1991) *Development* 112:231-240.
Liang & Richardson (1993) *J. Agric. Food Chem.* 41:1800-1807.
Marcotte *et al.* (1989) *Plant Cell* 1:969.
Millar *et al.* (1995a) *Science* 267(5201):1163-1166.
Millar *et al.* (1995b) *Science* 267(5201):1161-1163.
Millar *et al.* (1992) *Plant Cell* 4:1075-1087.
Millar and Kay (1991) *Plant Cell* 3:541-550.
Mozo *et al.* (1998) *Mo. Gen. Genet.* 258(5):562-570.
Murfet (1985) *Pisum sativum.* In *Handbook of Flowering Plants Vol. IV*, ed. A.H. Halevy. (CRC Press: Boca Raton, Florida), pp. 97-126.
Needleman and Wunsch (1970). *J. Mol. Biol.* 48: 443.
Newman *et al.* (1994) *Plant Physiol.* 106(4):1241-1255.
Odel *et al.* (1985) *Nature* 313:810.
Opperman *et al.* (1993) *Science* 263:221-223.
Pearson and Lipman (1988). *Proc. Natl. Acad. Sci. USA* 85: 2444.
Pearson *et al.* (1994). *Methods in Molecular Biology* 24, 307-31.
Pouwels *et al.* (1987) *Cloning Vectors: A Laboratory Manual*, 1985, supp.
Roshal *et al.* (1987) *EMBO J.* 6:1155.

Sambrook *et al.* (1989) In *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York.
Schaffer *et al.* (1998) *Cell* 93:1219-1229.
Schaffner & Sheen (1991) *Plant Cell* 3:997.
Schernthaner *et al.* (1988) *EMBO J.* 7:1249.
Shannon and Meeks-Wagner (1991) *Plant Cell* 3:877-892.
Siebertz *et al.* (1989) *Plant Cell* 1:961.
Smith *et al.* (1985) *Science* 229:1219-1224.
Smith and Waterman (1981). *Adv. Appl. Math.* 2: 482.
Stockhause *et al.* (1997) *The Plant Cell* 9:479-489.
Terada & Shimamoto (1990) *Mol. Gen. Genet.* 220:389.
Tijssen (1993). *Laboratory Techniques in Biochemistry and Molecular Biology -- Hybridization with Nucleic Acid Probes* Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York.
Wang & Tobin (1998) *Cell* 93:1207-1217.
Weissbach & Weissbach (1989) *Methods for Plant Molecular Biology*, Academic Press.
Zagotta *et al.* (1992) *Aust. J. Plant Physiol.* 19:411-418.
Zagotta *et al.* (1996) *Plant J.* 10(4):691-702.